

OPTIMAL PARAMETERS AND REGULARIZERS FOR IMAGE
RECONSTRUCTION PROBLEMS
- ONE DIMENSIONAL SETTING

ELISA DAVOLI

*Faculty of Mathematics, University of Vienna
Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria*

PAN LIU

*Department of Mathematics, Carnegie Mellon University
Forbes Avenue, Pittsburgh PA 15213, USA*

November 4, 2016

ABSTRACT. A new fractional order seminorm, $ICTV^r$, $r \in \mathbb{R}$, $r \geq 1$, is proposed in the one-dimensional setting, as a generalization of the standard $ICTV^k$ -seminorms, $k \in \mathbb{N}$. The fractional $ICTV^r$ -seminorms are shown to be intermediate between the standard $ICTV^k$ -seminorms of integer order. A bilevel learning scheme is proposed, where under a box constraint a simultaneous optimization with respect to the parameter α and the order r is performed. A numerical implementation of the learning scheme is provided, as well as an example where the optimal reconstruction is achieved for non-integer values of the order of derivation r .

1. INTRODUCTION

In the last decades Calculus of Variations and Partial Differential Equations (PDE) methods have proven to be very efficient in image denoising problems. Image denoising consists, roughly speaking, in recovering a noise-free clean picture u_c starting from a corrupted image $u_0 = u_c + \eta$, by filtering out the noise (or texture) encoded by η . One of the most successful variational approach to image denoising (see, for example [33, 34, 35]) relies on the ROF total variational functional

$$ROF(u) := \frac{1}{2} \int_I |u - u_0|^2 dx + \alpha TV(u), \quad (1.1)$$

introduced in [33]. Here $I = (0, 1)$ represents the domain of a one-dimensional image (a signal), $\alpha \in \mathbb{R}^+$, and $TV(u) = |u'|_{\mathcal{M}_b(I)}$, stands for the total variation of the measure u' on I (see [2, Definition 1.4]). An important role in determining the image reconstruction properties of the ROF functional is played by the parameter α . Indeed, if α is too large, then

E-mail addresses: elisa.davoli@univie.ac.at, panl@andrew.cmu.edu.

2010 Mathematics Subject Classification. 26B30, 94A08, 47J20.

Key words and phrases. infimal convolution, fractional derivative, optimization and control, computer vision and pattern recognition.

the total variation of u is too penalized and the image turns out to be over-smoothed, with a resulting loss of information on the internal edges of the picture. Conversely, if α is too small then the noise remains un-removed. The choice of the “best” parameter α then becomes an interesting task. In [20] the authors proposed a training scheme (\mathcal{B}) relying on a bilevel learning optimization defined in machine learning, namely on a semi-supervised learning scheme that optimally adapts itself to the given “perfect data” (see [11, 12, 22, 23, 37, 38]). This learning scheme searches for the optimal α so that the recovered image u_α , obtained as a minimizer of (1.1), optimizes the L^2 -distance from the clean image u_c . An implementation of (\mathcal{B}) equipped with total variation is the following:

Level 1.

$$\alpha_m := \arg \min_{\alpha > 0} \frac{1}{2} \int_I |u_\alpha - u_c|^2 dx,$$

Level 2.

$$u_\alpha := \arg \min_{u \in SBV(Q)} \left\{ \frac{1}{2} \int_I |u - u_0|^2 dx + \alpha TV(u) \right\}. \quad (1.2)$$

It is well known that the ROF model in (1.1) suffers drawbacks like the staircasing effect, and the training scheme (\mathcal{B}) inherits that feature, namely the optimized reconstruction function u_{α_m} also exhibits the staircasing effect. One approach to counteract this problem is to insert higher level derivatives in the regularizer (see [7, 9, 14, 31]). Two of the most successful image-reconstruction functionals among those involving mixed first and higher order terms are the infimal-convolution total variation ($ICTV$) [7] and the total generalized variation (TGV) models [31]. Note that they coincide with each other in the one-dimensional setting.

For $I := (0, 1) \subset \mathbb{R}$, $u \in BV(I)$, the $ICTV_\alpha^k$ regularizer is defined, for $k \in \mathbb{N}$, and $\alpha = (\alpha_0, \dots, \alpha_k) \in \mathbb{R}_+^{k+1}$, as

$$\begin{aligned} |u|_{ICTV_\alpha^{k+1}(I)} := \inf \bigg\{ & \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \alpha_1 |v'_0 - v_1|_{\mathcal{M}_b(I)} + \\ & \dots + \alpha_{k-1} |v'_{k-2} - v_{k-1}|_{\mathcal{M}_b(I)} + \alpha_k |v'_{k-1}|_{\mathcal{M}_b(I)} : \\ & v_i \in BV(I) \text{ for } 0 \leq i \leq k-1 \bigg\}. \end{aligned}$$

For instance, for $k = 1$, the $ICTV_{\alpha_0, \alpha_1}^2$ regularizer reads as

$$|u|_{ICTV_{\alpha_0, \alpha_1}^2(I)} := \inf \left\{ \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \alpha_1 |v'_0|_{\mathcal{M}_b(I)}, v_0 \in BV(I) \right\}.$$

Substituting $ICTV_{\alpha_0, \alpha_1}^2$ into (1.2) provides a bilevel learning scheme with $ICTV$ image reconstruction model. We recall that large values of α_1 will yield regularized solutions that are close to TV regularized reconstructions, and large values of α_0 will result in TV^2 type solutions (see, e.g., [32]). The best choice of parameters α_0 and α_1 is determined by an adaptation of the learning scheme (\mathcal{B}) above (see [19] for a detailed study).

In the existing literature a regularizer is fixed a priori, and the biggest effort is concentrated on studying how to identify the best parameters. In the case of the $ICTV^k$ model, this amounts to set manually the value of k first, and then determine the optimal α in (1.2). However, there is no evidence suggesting that $ICTV^2$ will always perform better than TV . In addition, the higher order seminorms $ICTV^k$, $k \geq 2$, have rarely been analyzed, and hence their performance is largely unknown. Numerical simulations show that for different images (signals in 1D), different orders of $ICTV^k$ might give different results. The main focus of this paper is exactly to investigate how to optimally tune both the weight α and the order k of the $ICTV_\alpha^k$ -seminorm, in order to achieve the best reconstructed image.

In this work we develop a bilevel learning scheme, not only for parameter learning, but also for determining the optimal order k of the regularizer $ICTV^k$ for image reconstruction. A straightforward modification of (\mathcal{B}) would be to just insert the order of the regularizer inside the learning level 2 in (1.2). Namely,

Level 1.

$$(\bar{\alpha}, \bar{k}) := \arg \min_{\alpha > 0, k \in \mathbb{N}} \frac{1}{2} \int_I |u_{\alpha, k} - u_c|^2 dx,$$

Level 2.

$$u_{\alpha, k} := \arg \min_{u \in SBV(Q)} \left\{ \frac{1}{2} \int_I |u - u_0|^2 dx + |u|_{ICTV_{\alpha}^{k+1}(I)} \right\}.$$

Often, in order to show the existence of a solution of the learning scheme and also for the numerical realization of the model, a *box constraint*

$$(\alpha, k) \in [A, B]^{k+1} \times [1, R], \quad (1.3)$$

where $A > 0$, $B > 0$, and $R > 1$ are fixed real numbers, needs to be imposed (see, e.g. [3, 18]). However, such constraint makes the above learning scheme less interesting. To be precise, restricting the analysis to the case in which $k \in \mathbb{N}$ is an integer, the box constraint (1.3) would only allow k to take finitely many values, and hence the optimized order \bar{k} of regularizer would simply be determined by performing scheme (\mathcal{B}) finitely many times, at each time with different values of k . In addition, finer texture effects, for which an “intermediate” reconstruction between the one provided by $ICTV^k$ and $ICTV^{k+1}$ for some $k \in \mathbb{N}$ would be needed, might be neglected in the optimization procedure.

Therefore, a main challenge in the setup of such a learning scheme is to give a meaningful interpolation between the spaces $ICTV^{k+1}$ and $ICTV^k$, and hence to guarantee that the collection of such spaces itself exhibits certain compactness and lower semicontinuity properties. For this purpose, we modify the definition of the $ICTV^k$ functionals by incorporating the theory of fractional Sobolev spaces, and we introduce the notion of fractional order $ICTV^{k+s}$ space (see Definition 3.1), where $k \in \mathbb{N}$ and $0 < s < 1$. For $k = 1$, our definition reads as follows.

$$|u|_{ICTV_{\alpha}^{1+s}(I)} := \inf \left\{ \alpha_0 |u' - sv_0|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_0|_{W^{s, 1+s(1-s)}(I)} : v_0 \in W^{s, 1+s(1-s)}(I), \int_I v_0(x) dx = 0 \right\}.$$

In addition, for every $k \in \mathbb{N}$ and $s \in [0, 1]$ we introduce the classes of functions with bounded infimal-convolution total variation seminorm

$$BCV_{\alpha}^{k+s}(I) := \left\{ u \in L^1(I) : |u|_{ICTV_{\alpha}^{k+s}(I)} < +\infty \right\}.$$

In the expressions above, $W^{s, 1+s(1-s)}(I)$ is the *fractional Sobolev space* of order s and integrability $1 + s(1 - s)$. In Theorem 3.2 we show that the $ICTV^{1+s}$ seminorm is indeed intermediate between $ICTV^1$ and $ICTV^2$, i.e., we prove that, up to subsequences,

$$\lim_{s \rightarrow 1} |u|_{ICTV_{\alpha}^{1+s}(I)} \geq |u|_{ICTV_{\alpha}^2(I)} \quad \text{and} \quad \lim_{s \rightarrow 0} |u|_{ICTV_{\alpha}^{1+s}(I)} = |u'|_{\mathcal{M}_b(I)}.$$

Namely, for $s \rightarrow 1$, the behavior of the $ICTV_\alpha^{1+s}$ -seminorm is close to the one of the standard $ICTV_\alpha^2$ -seminorm, whereas for $s \rightarrow 0$ it approaches the TV functional. We additionally prove (see Corollary 3.5) that analogous results hold for higher order $ICTV_\alpha^{k+s}$ -seminorms.

The advantage in working with such interpolation spaces is twofold. First, $ICTV^{k+s}$ is expected to inherit the advantages of fractional order derivatives, which have shown to be able to reduce the staircasing and contrast effects in noise removal problems (see, e.g. [10]). Second, they allow us to introduce the following improved learning scheme (\mathcal{R}) , which, under (1.3), simultaneously optimizes both with respect to the parameter α and to the order r of derivation.

Level 1.

$$(\bar{\alpha}, \bar{r}) := \arg \min \left\{ \int_I |u_{\alpha,r} - u_c|^2 dx, (\alpha, r) \in [A, B]^{\lfloor r \rfloor + 1} \times [1, R] \right\}, \quad (1.4)$$

Level 2.

$$u_{\alpha,r} := \arg \min_{u \in BCTV_\alpha^r(I)} \left\{ \int_I |u - u_0|^2 dx + ICTV_\alpha^r(u) \right\}. \quad (1.5)$$

In the definition above, $\lfloor r \rfloor$ denotes the largest integer smaller than or equal to r . Note that, according to the test noise-free image u_c , the level 1 (1.4) in our learning scheme (\mathcal{R}) directly indicates the higher order regularizer providing the best image reconstruction, as well as the associated corresponding optimal parameters.

In Section 5 a sketch of a numerical implementation of the scheme is proposed, as well as an explicit example for which the optimal reconstruction of a given signal with respect to an L^2 -error is achieved for non-integer values of r (see Figure 2). The two dimensional setting of fractional order $ICTV^s$ and TGV^s seminorm, as well as more extensive numerical analysis and examples for different type of images (images with large flat area and fine details, etc.) will be the subject of the follow-up work [17]. Ongoing research directions toward generalizations of the learning scheme (\mathcal{R}) are summarized in Section 6.

Our paper is organized as follows. In Section 2 we review the definitions and some basic properties of fractional order Sobolev spaces. In Section 3 we introduce the fractional order $ICTV^{k+s}$ seminorms and we study their main properties. In Section 4 we introduce our learning scheme (\mathcal{R}) . In particular, in Theorem 4.2 we show that (\mathcal{R}) admits a solution under the box constraint (4.1). In Section 5 we describe how to solve such minimizing problem (1.5) numerically and some experimental results are presented. Section 6 contains an outlook on some current work in progress aimed at providing an extension of the learning scheme (\mathcal{R}) to more general contexts.

2. THE THEORY OF FRACTIONAL SOBOLEV SPACES

In what follows we will assume that $I = (0, 1)$. We first recall a few results from the theory of Fractional Sobolev spaces. We refer to [21] for an introduction to the main results, and to [1, 26, 27, 30] and the references therein for a comprehensive treatment of the topic.

Definition 2.1 (Fractional Sobolev spaces). For $0 < s < 1$, $1 \leq p < +\infty$, and $u \in L^p(I)$, we define the *Gagliardo seminorm* of u by

$$|u|_{W^{s,p}(I)} := \left(\int_I \int_I \frac{|u(x) - u(y)|^p}{|x - y|^{1+sp}} dx dy \right)^{\frac{1}{p}}. \quad (2.1)$$

We say that $u \in W^{s,p}(I)$ if

$$\|u\|_{W^{s,p}(I)} := \|u\|_{L^p(I)} + |u|_{W^{s,p}(I)} < +\infty.$$

The following embedding results hold true ([21, Theorems 6.7, 6.10, and 8.2, and Corollary 7.2]).

Theorem 2.2 (Sobolev Embeddings - 1). *Let $s \in (0, 1)$ be given.*

1. *Let $p < \frac{1}{s}$. Then there exists a positive constant $C = C(p, s)$ such that for every $u \in W^{s,p}(I)$ there holds*

$$\|u\|_{L^q(I)} \leq C \|u\|_{W^{s,p}(I)} \quad (2.2)$$

for every $q \in [1, \frac{p}{1-sp}]$. If $q < \frac{p}{1-sp}$, then the embedding of $W^{s,1}(I)$ into $L^q(I)$ is also compact.

2. *Let $p = \frac{1}{s}$. Then the embedding in (2.2) holds for every $q \in [1, +\infty)$.*
3. *Let $p > \frac{1}{s}$. Then there exists a positive constant $C = C(p, s)$ such that for every $u \in W^{s,p}(I)$ we have*

$$\|u\|_{C^{0,\alpha}(I)} \leq C \|u\|_{W^{s,p}(I)},$$

with $\alpha := \frac{sp-1}{p}$.

The additional embedding result below is proved in [36, Corollary 19].

Theorem 2.3 (Sobolev Embeddings - 2). *Let $s \geq r$, $p \leq q$ and $s - 1/p \geq r - 1/q$, with $0 < r \leq s < 1$, and $1 \leq p \leq q \leq +\infty$. Then*

$$W^{s,p}(I) \subset W^{r,q}(I)$$

and

$$|u|_{W^{r,q}(I)} \leq \frac{36}{rs} |u|_{W^{s,p}(I)}$$

The next inequality is a special case of [4, Theorem 1] and [29, Theorem 1].

Theorem 2.4 (Poincaré Inequality). *Let $p \geq 1$, and let $sp < 1$. There exists a constant $C > 0$ such that*

$$\left\| u - \int_I u(x) dx \right\|_{L^{\frac{p}{1-sp}}(I)}^p \leq \frac{Cs(1-s)}{(1-sp)^{p-1}} |u|_{W^{s,p}(I)}^p.$$

It is possible to construct a continuous extension operator from $W^{s,1}(I)$ to $W^{s,1}(\mathbb{R})$ (see, e.g., [21, Theorem 5.4]).

Theorem 2.5 (Extension Operator). *Let $s \in (0, 1)$, and let $1 \leq p < +\infty$. Then $W^{s,p}(I)$ is continuously embedded in $W^{s,p}(\mathbb{R})$, namely there exists a constant $C = C(p, s)$ such that for every $u \in W^{s,p}(I)$ there exists $\tilde{u} \in W^{s,p}(\mathbb{R})$ satisfying $\tilde{u}|_I = u$ and*

$$\|\tilde{u}\|_{W^{s,p}(\mathbb{R})} \leq C \|u\|_{W^{s,p}(I)}.$$

The next two theorems ([39, Section 2.2.2, Remark 3, and Section 2.11.2]) yield an identification between fractional Sobolev spaces and Besov spaces in \mathbb{R} , and guarantee the reflexivity of Besov spaces $B_{p,q}^s$ for p, q finite.

Theorem 2.6 (Identification with Besov spaces). *If $1 \leq p < +\infty$ and $s \in \mathbb{R}^+ \setminus \mathbb{N}$, then*

$$W^{s,p}(\mathbb{R}) = B_{p,p}^s(\mathbb{R})$$

Theorem 2.7 (Reflexivity of Besov spaces). *Let $-\infty < s < +\infty$, $1 \leq p < +\infty$ and $0 < q < +\infty$. Then*

$$(B_{p,q}^s(\mathbb{R}))' = B_{p',q'}^{-s}(\mathbb{R}),$$

where $(B_{p,q}^s(\mathbb{R}))'$ is the dual of the Besov space $B_{p,q}^s(\mathbb{R})$, and where p' and q' are the conjugate exponent of p and q , respectively.

In view of Theorems 2.6 and 2.7 the following characterization holds true.

Corollary 2.8 (Reflexivity of Fractional Sobolev spaces). *Let $1 < p < +\infty$ and $s \in \mathbb{R}^+ \setminus \mathbb{N}$. Then the fractional Sobolev space $W^{s,p}(\mathbb{R})$ is reflexive.*

We conclude this section by recalling two theorems describing the limit behavior of the Gagliardo seminorm as $s \rightarrow 1$ and $s \rightarrow 0$, respectively. The first result has been proved in [5, Theorem 3 and Remark 1], and [15, Theorem 1].

Theorem 2.9 (Asymptotic behavior as $s \rightarrow 1$). *Let $u \in BV(I)$. Then*

$$\lim_{s \rightarrow 1} (1-s) |u|_{W^{s,1}(I)} = |u'|_{\mathcal{M}_b(I)}.$$

Similarly, the asymptotic behavior of the Gagliardo seminorm has been characterized as $s \rightarrow 0$ in [29, Theorem 3].

Theorem 2.10 (Asymptotic behavior as $s \rightarrow 0$). *Let $u \in \cup_{0 < s < 1} W^{s,1}(\mathbb{R})$. Then,*

$$\lim_{s \rightarrow 0} s |u|_{W^{s,1}(\mathbb{R})} = \|u\|_{L^1(\mathbb{R})}.$$

3. THE FRACTIONAL ICTV SEMINORM

In this section we define the fractional *ICTV* seminorm, and we prove some first properties.

Definition 3.1 (The Fractional *ICTV* Space). *Let $0 < s < 1$, $k \in \mathbb{N}$, and let $\alpha = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k) \in \mathbb{R}_+^{k+1}$. For every $u \in L^1(I)$, we define its fractional $ICTV^{k+s}$ seminorm as follows.*

For $k = 1$ we set

$$|u|_{ICTV_\alpha^{1+s}(I)} := \inf \left\{ \alpha_0 |u' - sv_0|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_0|_{W^{s,1+s(1-s)}(I)} : \right. \\ \left. v_0 \in W^{s,1+s(1-s)}(I), \int_I v_0(x) dx = 0 \right\}.$$

For $k > 1$ we define

$$|u|_{ICTV_\alpha^{k+s}(I)} := \inf \left\{ \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \alpha_1 |v'_0 - v_1|_{\mathcal{M}_b(I)} + \right. \\ \left. \dots + \alpha_{k-1} |v'_{k-2} - sv_{k-1}|_{\mathcal{M}_b(I)} + \alpha_k s(1-s) |v_{k-1}|_{W^{s,1+s(1-s)}(I)} : \right. \\ \left. v_i \in BV(I) \text{ for } 0 \leq i \leq k-2, v_{k-1} \in W^{s,1+s(1-s)}(I), \int_I v_{k-1}(x) dx = 0. \right\}$$

For $0 \leq s < 1$, $k \in \mathbb{N}$, $\alpha = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k) \in \mathbb{R}_+^{k+1}$, we say that $u \in BCV_\alpha^{k+s}(I)$ if

$$\|u\|_{BCV_\alpha^{k+s}(I)} := \|u\|_{L^1(I)} + |u|_{ICTV_\alpha^{k+s}(I)} < +\infty,$$

and we write $u \in BCV^{k+s}(I)$ if there exists $\alpha \in \mathbb{R}_+^{k+1}$ such that $u \in BCV_\alpha^{k+s}(I)$. Note that if $u \in BCV_\alpha^{k+s}(I)$ for some $\alpha \in \mathbb{R}_+^{k+1}$, then $u \in BCV_\beta^{k+s}(I)$ for every $\beta \in \mathbb{R}_+^{k+1}$.

We observe that the $ICTV^{k+s}$ seminorm is actually “intermediate” between the $ICTV^k$ seminorm and the $ICTV^{k+1}$ seminorm. To be precise, we have the following identification.

Theorem 3.2. *For every $u \in BV(I)$, up to the extraction of a (non-relabeled) subsequence there holds*

$$\liminf_{s \rightarrow 1} |u|_{ICTV_\alpha^{1+s}(I)} \geq |u|_{ICTV_\alpha^2(I)} \quad \text{and} \quad \lim_{s \rightarrow 0} |u|_{ICTV_\alpha^{1+s}(I)} = |u'|_{\mathcal{M}_b(I)}.$$

Before proving the theorem we state and prove an intermediate result that will be crucial in determining the asymptotic behavior of the $ICTV^{1+s}$ seminorm as $s \rightarrow 1$.

Proposition 3.3. *Let $u \in BV(I)$. Then*

$$\limsup_{s \rightarrow 1} (1-s) |u|_{W^{s,1+s(1-s)}(I)} \leq |u'|_{\mathcal{M}_b(I)}.$$

Proof. We first observe that for $x, y \in I$ there holds

$$|x - y|^{1+s+s^2(1-s)} \leq |x - y|^{1+s}.$$

Let now $u \in BV(I)$. Then

$$|u|_{W^{s,1+s(1-s)}(I)} \leq |u'|_{\mathcal{M}_b(I)}^{\frac{s(1-s)}{1+s(1-s)}} |u|_{W^{s,1}(I)}^{\frac{1}{1+s(1-s)}} \left(\int_I \int_I \frac{dx dy}{|x - y|^{s^2(1-s)}} \right)^{\frac{1}{1+s(1-s)}}.$$

Therefore,

$$\limsup_{s \rightarrow 1} (1-s) |u|_{W^{s,1+s(1-s)}(I)} \leq \limsup_{s \rightarrow 1} (1-s) |u|_{W^{s,1}(I)}^{\frac{1}{1+s(1-s)}}.$$

In view of Theorem 2.9 there holds

$$\lim_{s \rightarrow 1} |u|_{W^{s,1}(I)}^{\frac{s(1-s)}{1+s(1-s)}} = \lim_{s \rightarrow 1} e^{\frac{s(1-s)}{1+s(1-s)} \log |u|_{W^{s,1}(I)}} = 1.$$

Thus, again by Theorem 2.9 we conclude that

$$\limsup_{s \rightarrow 1} (1-s) |u|_{W^{s,1+s(1-s)}(I)} \leq \limsup_{s \rightarrow 1} (1-s) |u|_{W^{s,1}(I)} = |u'|_{\mathcal{M}_b(I)}.$$

□

A crucial ingredient in the proof of Theorem 3.2 is a compactness and lower semicontinuity result for maps with null averages and weighted $W^{s,1+s(1-s)}$ -seminorm.

Proposition 3.4. *Let $\{s_n\} \subset (0, 1)$ be such that $s_n \rightarrow \bar{s}$, with $\bar{s} \in (0, 1]$. For every $n \in \mathbb{N}$ let $v_n \in W^{s_n,1+s_n(1-s_n)}(I)$ satisfy $\int_I v_n(x) dx = 0$, and*

$$\sup_{n \geq 1} s_n(1-s_n) |v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} < +\infty. \quad (3.1)$$

Then, for $\bar{s} \in (0, 1)$, and up to the extraction of a (non-relabeled) subsequence, there exists $\bar{v} \in W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)$ such that

$$v_n \rightarrow \bar{v} \quad \text{strongly in } L^1(I),$$

and

$$\liminf_{n \rightarrow \infty} s_n(1-s_n) |v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} \geq \bar{s}(1-\bar{s}) |\bar{v}|_{W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)}. \quad (3.2)$$

The analogous statement holds for $\bar{s} = 1$, by replacing $W^{\bar{s},1+\bar{s}(1-\bar{s})}(I)$ with $BV(I)$, and (3.2) with

$$\liminf_{n \rightarrow \infty} s_n(1-s_n) |v_n|_{W^{s_n,1+s_n(1-s_n)}(I)} \geq |\bar{v}'|_{\mathcal{M}_b(I)}.$$

Proof. We first observe that for $x, y \in I$, and $s < t$, we have $|x - y|^{1+s} > |x - y|^{1+t}$. Hence, in view of (2.1) there holds

$$|u|_{W^{s,p}(I)} < |u|_{W^{t,p}(I)} \quad (3.3)$$

for $1 \leq p < +\infty$, and for every $u \in W^{t,p}(I)$.

Without loss of generality (and up to the extraction of a non-relabeled subsequence) we can assume that the sequences $\{s_n\}$ and $\{s_n(1 - s_n)\}$ converge monotonically to \bar{s} and $\bar{s}(1 - \bar{s})$, respectively. Therefore, according to the value of \bar{s} only 4 situations can arise:

Case 1: $0 < \bar{s} < \frac{1}{2}$: $s_n \searrow \bar{s}$ and $s_n(1 - s_n) \searrow \bar{s}(1 - \bar{s})$;

Case 2: $\frac{1}{2} \leq \bar{s} < 1$: $s_n \searrow \bar{s}$ and $s_n(1 - s_n) \nearrow \bar{s}(1 - \bar{s})$;

Case 3: $\frac{1}{2} < \bar{s} \leq 1$: $s_n \nearrow \bar{s}$ and $s_n(1 - s_n) \searrow \bar{s}(1 - \bar{s})$;

Case 4: $0 < \bar{s} \leq \frac{1}{2}$: $s_n \nearrow \bar{s}$ and $s_n(1 - s_n) \nearrow \bar{s}(1 - \bar{s})$.

We first consider Case 2. By (3.1) there exists a constant C such that

$$\sup_{n \geq 1} |v_n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \leq C.$$

We point out that the function $f : (0, 1) \rightarrow \mathbb{R}$, defined as

$$f(x) := x - \frac{1}{1 + x(1 - x)} \quad \text{for every } x \in (0, 1),$$

is strictly increasing on $(0, 1)$. Thus, we can apply Theorem 2.3 with $s = s_n$, $r = \bar{s}$, $p = 1 + s_n(1 - s_n)$, and $q = 1 + \bar{s}(1 - \bar{s})$ and we obtain

$$|v_n|_{W^{\bar{s}, 1+\bar{s}(1-\bar{s})}(I)} \leq C |v_n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \leq C. \quad (3.4)$$

Thus, by Theorem 2.4 and Corollary 2.8 there exists $\bar{v} \in W^{\bar{s}, 1+\bar{s}(1-\bar{s})}(I)$ such that, up to the extraction of a (non-relabeled) subsequence, we have

$$v_n \rightharpoonup \bar{v} \quad \text{weakly in } W^{\bar{s}, 1+\bar{s}(1-\bar{s})}(I).$$

By the lower semicontinuity of the $W^{\bar{s}, 1+\bar{s}(1-\bar{s})}(I)$ norm with respect to the weak convergence, and by (3.4) we deduce the inequality

$$\begin{aligned} \bar{s}(1 - \bar{s}) |\bar{v}|_{W^{\bar{s}, 1+\bar{s}(1-\bar{s})}(I)} &\leq \liminf_{n \rightarrow +\infty} \bar{s}(1 - \bar{s}) |v_n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \\ &= \liminf_{n \rightarrow +\infty} s_n(1 - s_n) |v_n|_{W^{s_n, 1+s_n(1-s_n)}(I)}. \end{aligned}$$

In Case 1 we observe that the function $g : (0, 1) \rightarrow \mathbb{R}$, defined as

$$g(x) := \frac{1}{1 + x(1 - x)} \quad \text{for every } x \in (0, 1),$$

is strictly decreasing in $(0, \frac{1}{2}]$. Since $s_1 \geq s_n \geq \bar{s}$ for every n , there holds

$$\frac{1}{1 + s_1(1 - s_1)} \geq \frac{1}{1 + \bar{s}(1 - \bar{s})},$$

and by the properties of the functions f and g ,

$$s_n - \frac{1}{1 + s_n(1 - s_n)} \geq \bar{s} - \frac{1}{1 + \bar{s}(1 - \bar{s})} \geq \bar{s} - \frac{1}{1 + s_1(1 - s_1)}.$$

By (3.1) there exists a constant C such that

$$\sup_{n \geq 1} |v_n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \leq C.$$

Choosing $s = s_n$, $r = \bar{s}$, $p = 1 + s_n(1 - s_n)$, and $q = 1 + s_1(1 - s_1)$ in Theorem 2.3 we have

$$|v_n|_{W^{\bar{s}, 1+s_1(1-s_1)}(I)} \leq |v_n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \leq C.$$

Thus, by Theorem 2.4 there exists a map \bar{v} such that, up to the extraction of a (non-relabeled) subsequence, there holds

$$v_n \rightharpoonup \bar{v} \quad \text{weakly in } W^{\bar{s}, 1+s_1(1-s_1)}(I),$$

and by Theorem 2.2 also strongly in $L^1(I)$. In particular, Fatou's Lemma yields

$$|v|_{W^{\bar{s}, 1+\bar{s}(1-\bar{s})}(I)}^{1+\bar{s}(1-\bar{s})} \leq \liminf_{n_k \rightarrow +\infty} |v_{n_k}|_{W^{s_{n_k}, 1+s_{n_k}(1-s_{n_k})}(I)}^{1+s_{n_k}(1-s_{n_k})},$$

which in turn implies the thesis.

We omit the proof of the result in Case 4, and in Case 3 for $\bar{s} < 1$, as they follow from analogous arguments. Regarding Case 3 for $\bar{s} = 1$, by (3.1) and (3.3) there exists a constant C such that

$$(1 - s_n) |v_n|_{W^{s_n, 1}(I)} \leq C,$$

for every $n \in \mathbb{N}$. The thesis follows then by [5, Theorem 4]. \square

We now prove Theorem 3.2.

Proof of Theorem 3.2. Let $v_0 \in BV(I)$ be such that

$$|u|_{ICTV_\alpha^2(I)} = \alpha_0 |u' - v_0|_{\mathcal{M}_b(I)} + \alpha_1 |v'_0|_{\mathcal{M}_b(I)}.$$

In view of Proposition 3.3 there holds

$$\begin{aligned} \limsup_{s \rightarrow 1} |u|_{ICTV_\alpha^{1+s}(I)} &\leq \limsup_{s \rightarrow 1} \left\{ \alpha_0 \left| u' - sv_0 + s \int_I v_0(x) dx \right|_{\mathcal{M}_b(I)} \right. \\ &\quad \left. + \alpha_1 s(1-s) |v_0|_{W^{s, 1+s(1-s)}(I)} \right\} \\ &\leq |u|_{ICTV_\alpha^2(I)} + \left| \int_I v_0(x) dx \right|. \end{aligned} \quad (3.5)$$

For every $s \in (0, 1)$, let $v_0^s \in W^{s, 1+s(1-s)}(I)$ be such that $\int_I v_0^s(x) dx = 0$, and

$$\alpha_0 |u' - sv_0^s|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_0^s|_{W^{s, 1+s(1-s)}(I)} \leq |u|_{ICTV_\alpha^{1+s}(I)} + (1-s). \quad (3.6)$$

In view of (3.5) and Proposition 3.4, there exists $v \in BV(I)$ such that, up to the extraction of a (non-relabeled) subsequence,

$$v_0^s \rightarrow v \quad \text{strongly in } L^1(I),$$

and

$$\lim_{s \rightarrow 1} s(1-s) |v_0^s|_{W^{s, 1+s(1-s)}(I)} \geq |v'|_{BV(I)}.$$

Passing to the limit in (3.6) we deduce the inequality

$$|u|_{ICTV_\alpha^2(I)} \leq \alpha_0 |u' - v|_{\mathcal{M}_b(I)} + \alpha_1 |v'|_{BV(I)} \leq \liminf_{s \rightarrow 1} |u|_{ICTV_\alpha^{1+s}(I)},$$

which in turn implies the thesis.

To study the case $s \rightarrow 0$, we first observe that

$$\sup_{s \in (0, 1)} |u|_{ICTV_\alpha^{1+s}(I)} \leq |u'|_{BV(I)}. \quad (3.7)$$

Thus we only need to prove the opposite inequality. To this aim, for every $s \in (0, 1)$ let $v_0^s \in W^{s, 1+s(1-s)}(I)$ be such that $\int_I v_0^s(x) dx = 0$, and

$$\alpha_0 |u' - sv_0^s|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_0^s|_{W^{s, 1+s(1-s)}(I)} \leq |u|_{ICTV_\alpha^{1+s}(I)} + s. \quad (3.8)$$

Since $s(1+s(1-s)) < 1$ for $s \in (0, 1)$, by (3.7) and in view of Theorem 2.4 there holds

$$sv_0^s \rightarrow 0 \quad \text{strongly in } L^1(I).$$

Passing to the limit in (3.8) we deduce the inequality

$$|u'|_{\mathcal{M}_b(I)} \leq \limsup_{s \rightarrow 0} |u|_{ICTV_\alpha^{1+s}(I)}.$$

The thesis follows owing to (3.7). \square

Corollary 3.5. *Let $k \geq 2$. For every $u \in BV(I)$, up to the extraction of a (non-relabelled) subsequence there holds*

$$\liminf_{s \rightarrow 1} |u|_{ICTV_\alpha^{k+s}(I)} \geq |u|_{ICTV_\alpha^{k+1}(I)} \quad \text{and} \quad \lim_{s \rightarrow 0} |u|_{ICTV_\alpha^{k+s}(I)} = |u|_{ICTV_\alpha^k(I)},$$

where $\hat{\alpha} := (\alpha_0, \dots, \alpha_{k-1}) \in \mathbb{R}_+^k$.

Proof. The result follows by straightforward adaptations of the arguments in the proof of Theorem 3.2. \square

We proceed by showing that the minimization problem in Definition 3.1 has a solution.

Proposition 3.6. *If the infimum in Definition 3.1 is finite, then it is attained.*

Proof. Let $k = 1$. Let $\alpha \in \mathbb{R}_+^2$, and let $u \in BCV_\alpha^{k+s}(I)$. We need to show that

$$|u|_{ICTV_\alpha^{1+s}(I)} = \min \left\{ \alpha_0 |u' - sv|_{\mathcal{M}_b(I)} + s(1-s)\alpha_1 |v|_{W^{s, 1+s(1-s)}(I)} : \right. \\ \left. v \in W^{s, 1+s(1-s)}(I), \int_I v(x) dx = 0 \right\}. \quad (3.9)$$

We first observe that $u \in BV(I)$.

Indeed, let $\eta > 0$, and let $v \in W^{s, 1+s(1-s)}(I)$ be such that $\int_I v(x) dx = 0$, and

$$\alpha_0 |u' - sv|_{\mathcal{M}_b(I)} + s(1-s)\alpha_1 |v|_{W^{s, 1+s(1-s)}(I)} \leq |u|_{ICTV_\alpha^{1+s}(I)} + \eta.$$

By Hölder inequality there holds

$$\begin{aligned} \alpha_0 |u'|_{\mathcal{M}_b(I)} &\leq \alpha_0 |u' - sv|_{\mathcal{M}_b(I)} + s\alpha_0 \|v\|_{L^1(I)} \\ &\leq \alpha_0 |u' - sv|_{\mathcal{M}_b(I)} + s\alpha_0 |v|_{W^{s, 1+s(1-s)}(I)} + s\alpha_0 \|v\|_{L^{1+s(1-s)}(I)} \\ &\leq |u|_{ICTV_\alpha^{1+s}(I)} + \eta + s(\alpha_0 - (1-s)\alpha_1) |v|_{W^{s, 1+s(1-s)}(I)} + s\alpha_0 \|v\|_{L^{1+s(1-s)}(I)}. \end{aligned}$$

Let now $\{v_n\} \subset W^{s, 1+s(1-s)}(I)$ be a minimizing sequence for (3.9). Since $s(1+s(1-s)) < 1$ for $s \in (0, 1)$, by Theorem 2.2 there exists a constant C such that

$$\sup_{n \in \mathbb{N}} \|v_n\|_{W^{s, 1+s(1-s)}(I)} \leq C.$$

Thus, by Corollary 2.8 there exists $\bar{v} \in W^{s, 1+s(1-s)}(I)$ such that, up to the extraction of a (non-relabelled) subsequence, there holds

$$v_n \rightharpoonup \bar{v} \quad \text{weakly in } W^{s, 1+s(1-s)}(I),$$

and hence by Theorem 2.2,

$$v_n \rightarrow \bar{v} \quad \text{strongly in } L^1(I).$$

The thesis follows now by the lower semicontinuity of the total variation and the $W^{s,1+s(1-s)}$ norm with respect to the L^1 convergence and the weak convergence in $W^{s,1+s(1-s)}(I)$, respectively.

For $k = 2$, let $\{v_0^n\} \subset BV(I)$ and $\{v_1^n\} \subset W^{s,1+s(1-s)}(I)$ with $\int_I v_1^n(x) dx = 0$ for every $n \in \mathbb{N}$ be such that

$$\begin{aligned} & \lim_{n \rightarrow +\infty} \{ \alpha_0 |u' - v_0^n|_{\mathcal{M}_b(I)} + \alpha_1 |(v_0^n)' - s v_1^n|_{\mathcal{M}_b(I)} + \alpha_2 s(1-s) |v_1^n|_{W^{1,s(1-s)}(I)} \} \\ & = ICTV_\alpha^{2+s}(I). \end{aligned}$$

Since $s(1+s(1-s)) < 1$ for $s \in (0,1)$, by Theorem 2.2 we obtain that $\{v_1^n\}$ is uniformly bounded in $W^{s,1+s(1-s)}(I)$. Therefore, $\{v_0^n\}$ is uniformly bounded in $BV(I)$, and there exist $v_0 \in BV(I)$ and $v_1 \in W^{s,1+s(1-s)}(I)$, with $\int_I v_1(x) dx = 0$, such that, up to the extraction of a (non-relabeled) subsequence,

$$v_0^n \rightharpoonup^* v_0 \quad \text{weakly}^* \text{ in } BV(I),$$

and

$$v_1^n \rightharpoonup v_1 \quad \text{weakly in } W^{s,1+s(1-s)}(I).$$

In particular, by Theorem 2.2,

$$v_1^n \rightarrow v_1 \quad \text{strongly in } L^1(I).$$

The minimality of v_0 and v_1 is a consequence of lower semicontinuity. The thesis for $k > 2$ follows by analogous arguments. \square

We observe that the $ICTV^{k+s}$ seminorms are all equivalent to the total variation seminorm.

Lemma 3.7. *For every $k \geq 1$ and $0 < s < 1$, we have*

$$BV(I) \sim BCV^k(I) \sim BCV^{k+s}(I) \sim BCV^{k+1}(I).$$

Proof. We only show that

$$BV(I) \sim BCV^{1+s}(I) \sim BCV^2(I). \quad (3.10)$$

The proof of the inequality for $k > 1$ is analogous. In view of (3.7), to prove the first equivalence relation in (3.10) we only need to show that there exist a constant C and a multi-index $\alpha \in \mathbb{R}_+^2$ such that

$$|u'|_{\mathcal{M}_b(I)} \leq C |u|_{ICTV_\alpha^{1+s}(I)}.$$

By Theorem 2.2 we have

$$\begin{aligned} |u'|_{\mathcal{M}_b(I)} & \leq |u' - s v_0|_{\mathcal{M}_b(I)} + s |v_0|_{L^1(I)} \\ & \leq |u' - s v_0|_{\mathcal{M}_b(I)} + C s |v_0|_{W^{s,1+s(1-s)}(I)} \\ & = |u' - s v_0|_{\mathcal{M}_b(I)} + \frac{C}{(1-s)} s(1-s) |v_0|_{W^{s,1+s(1-s)}(I)} \end{aligned}$$

for every $v_0 \in W^{s,1+s(1-s)}(I)$. Thus

$$|u'|_{\mathcal{M}_b(I)} \leq C |u|_{ICTV_{1, \frac{C}{1-s}}^{1+s}(I)}$$

for every $s \in (0,1)$. This completes the proof of the first equivalence in (3.10). Property (3.10) follows now by [6, Theorem 3.3]. \square

We conclude this section with a proposition that will be crucial in establishing our new learning scheme.

Proposition 3.8. *Let $k \in \mathbb{N}$, $\{s_n\} \subset (0, 1)$, and $\{\alpha^n\} \subset \mathbb{R}_+^{k+1}$, with $\alpha^n = (\alpha_0^n, \alpha_1^n, \dots, \alpha_k^n)$ satisfying*

$$0 < A < \inf \{\alpha_i^n, n \geq 1, 0 \leq i \leq k\} \leq \sup \{\alpha_i^n, n \geq 1, 0 \leq i \leq k\} < B < +\infty. \quad (3.11)$$

Let $u_n \in BC V_{\alpha^n}^{k+s_n}(I)$ be such that

$$\sup_{n \in \mathbb{N}} \left\{ \|u_n\|_{BC V_{\alpha^n}^{k+s_n}(I)} \right\} < +\infty. \quad (3.12)$$

Then, up to the extraction of a (non-relabeled) subsequence, there exist $\bar{s} \in [0, 1]$, $\alpha \in \mathbb{R}_+^{k+1}$ and $u \in BC V_{\alpha}^{k+\bar{s}}(I)$ such that $s_n \rightarrow \bar{s}$, $\alpha^n \rightarrow \alpha$, and

$$u_n \xrightarrow{*} u \text{ in } BV(I). \quad (3.13)$$

In addition, if $\bar{s} \in (0, 1]$ there holds

$$|u|_{ICTV_{\alpha}^{k+\bar{s}}(I)} \leq \liminf_{n \rightarrow +\infty} |u_n|_{ICTV_{\alpha^n}^{k+s_n}(I)}. \quad (3.14)$$

If $\bar{s} = 0$ we have

$$|u|_{ICTV_{\hat{\alpha}}^k(I)} \leq \liminf_{n \rightarrow +\infty} |u_n|_{ICTV_{\alpha^n}^{k+s_n}(I)},$$

where $\hat{\alpha} \in \mathbb{R}_+^k$ is the multi-index $\hat{\alpha} := (\alpha_0, \dots, \alpha_{k-1})$, and

$$\alpha_0 |u'|_{\mathcal{M}_b(I)} \leq \liminf_{n \rightarrow +\infty} |u_n|_{ICTV_{\alpha^n}^{k+s_n}(I)}, \quad (3.15)$$

for $k > 1$ and $k = 1$, respectively.

Proof. We prove the statement for $k = 1$. The proof of the result for $k > 1$ follows via straightforward modifications. For $k = 1$, we have $\{\alpha^n\} \subset \mathbb{R}^2$, and by (3.11) up to the extraction of a (non-relabeled) subsequence there holds

$$(\alpha_0^n, \alpha_1^n) \rightarrow (\alpha_0, \alpha_1), \text{ where } A < \alpha_0, \alpha_1 < B. \quad (3.16)$$

In addition, since $\{s_n\}$ is bounded, there exists $\bar{s} \in [0, 1]$ such that $s_n \rightarrow \bar{s}$. By Proposition 3.6 we deduce that there exists $v_0^n \in W^{s_n, 1+s_n(1-s_n)}(I)$ such that

$$|u|_{ICTV_{\alpha^n}^{1+s_n}(I)} = \alpha_0^n |u'_n - s_n v_0^n|_{\mathcal{M}_b(I)} + \alpha_1^n s_n (1 - s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)}.$$

Thus, by Theorem 2.4 and since $\int_I v_0^n(x) dx = 0$, we have (note that $s_n < 1$)

$$\begin{aligned} |u'_n|_{\mathcal{M}_b(I)} &\leq |u'_n - s_n v_0^n|_{\mathcal{M}_b(I)} + s_n |v_0^n|_{L^1(I)} \\ &\leq C \left\{ |u'_n - s_n v_0^n|_{\mathcal{M}_b(I)} + s_n (1 - s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \right\}. \end{aligned}$$

Hence, by (3.12) and (3.16),

$$\sup_{n \in \mathbb{N}} \left\{ \|u_n\|_{BV(I)} \right\} \leq C \sup_{n \in \mathbb{N}} \left\{ \|u_n\|_{BC V_{\alpha^n}^{1+s_n}(I)} \right\} < +\infty$$

which implies (3.13).

Note that for any $0 < s < 1$,

$$1 + s(1 - s) - \frac{1}{1 - s} = \frac{1}{1 - s} (-s + s(1 - s)^2) < 0.$$

Hence by applying again Theorem 2.4 we obtain

$$\|v_0^n\|_{L^{1+s_n(1-s_n)}(I)} \leq \|v_0^n\|_{L^{\frac{1}{1-s_n}}(I)} \leq C s_n (1 - s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)},$$

which by (3.11) implies

$$\begin{aligned} s_n(1-s_n) \|v_n\|_{W^{s_n, 1+s_n(1-s_n)}(I)} &\leq \|v_0^n\|_{L^{1+s_n(1-s_n)}} + s_n(1-s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \\ &\leq C s_n(1-s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \leq C \|u_n\|_{BCV_{\alpha^n}^{1+s_n}(I)}. \end{aligned}$$

Therefore, by (3.12) we deduce the uniform bound

$$\sup_{n \geq 1} s_n(1-s_n) \|v_n\|_{W^{s_n, 1+s_n(1-s_n)}(I)} < +\infty. \quad (3.17)$$

We subdivide the remaining part of the proof of Proposition 3.8 into 2 cases.

Case 1: $\bar{s} \in (0, 1]$.

Assume first that $\bar{s} < 1$. By Proposition 3.4 there exists $v_0 \in W^{\bar{s}, 1+\bar{s}(1-\bar{s})}(I)$ such that

$$v_0^n \rightarrow v_0 \quad \text{strongly in } L^1(I), \quad (3.18)$$

and

$$\liminf_{n \rightarrow \infty} s_n(1-s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \geq \bar{s}(1-\bar{s}) |v_0|_{W^{\bar{s}, 1+\bar{s}(1-\bar{s})}(I)}. \quad (3.19)$$

By (3.18) and since $0 < \bar{s} < 1$, there holds

$$\begin{aligned} &\liminf_{n \rightarrow \infty} |u_n|_{ICTV_{\alpha^n}^{1+s_n}(I)} \\ &\geq \liminf_{n \rightarrow \infty} \alpha_0^n |u'_n - s_n v_0^n|_{\mathcal{M}_b(I)} + \liminf_{n \rightarrow \infty} \alpha_1^n s_n(1-s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \\ &\geq \alpha_0 |u' - \bar{s} v_0|_{\mathcal{M}_b(I)} + \alpha_1 \bar{s}(1-\bar{s}) |v_0|_{W^{\bar{s}, 1+\bar{s}(1-\bar{s})}(I)} \geq |u|_{ICTV_{\alpha}^{1+\bar{s}}(I)}, \end{aligned} \quad (3.20)$$

where in the last inequality we used the definition of the $ICTV_{\alpha}^{1+\bar{s}}$ -seminorm. In particular, $u \in BCV_{\alpha}^{1+\bar{s}}(I)$, and (3.14) is satisfied. The proof of the proposition for $\bar{s} = 1$ follows via the same argument, and by replacing $\bar{s}(1-\bar{s}) |v_0|_{W^{\bar{s}, 1+\bar{s}(1-\bar{s})}(I)}$ in (3.19) and (3.20) with $|v'_0|_{\mathcal{M}_b(I)}$.

Case 2: $\bar{s} = 0$. In view of (3.17) and Theorem 2.4, up to the extraction of a (non-relabelled) subsequence we deduce that

$$s_n v_n \rightarrow 0 \quad \text{strongly in } L^1(I).$$

Hence, there holds

$$\begin{aligned} &\liminf_{n \rightarrow \infty} |u_n|_{ICTV_{\alpha^n}^{1+s_n}(I)} \\ &\geq \liminf_{n \rightarrow \infty} \alpha_0^n |u'_n - s_n v_0^n|_{\mathcal{M}_b(I)} + \liminf_{n \rightarrow \infty} \alpha_1^n s_n(1-s_n) |v_0^n|_{W^{s_n, 1+s_n(1-s_n)}(I)} \\ &\geq \alpha_0 |u'|_{\mathcal{M}_b(I)}, \end{aligned}$$

which in turn implies (3.15). This completes the proof of the proposition. \square

4. THE BILEVEL LEARNING SCHEME WITH RESPECT TO PARAMETER AND REGULARIZER

Let $r \geq 1$ be given and let $\lfloor r \rfloor$ denote the largest integer smaller than or equal to r . We propose the following learning scheme (\mathcal{R}) which takes into account the order r of the regularizer and the parameter $\alpha \in \mathbb{R}_+^{\lfloor r \rfloor + 1}$ simultaneously. We restrict our analysis to the case in which α and r satisfy the *box constraint*

$$(\alpha, r) \in [A, B]^{\lfloor r \rfloor + 1} \times [1, R] \quad (4.1)$$

where $A > 0$, $B > 0$, and $R > 1$ are fixed real numbers.

Our new learning scheme (\mathcal{R}) is defined as follows:

Level 1.

$$(\bar{\alpha}, \bar{r}) := \arg \min \left\{ \int_I |u_{\alpha, r} - u_c|^2 dx, (\alpha, r) \in [A, B]^{\lfloor r \rfloor + 1} \times [1, R] \right\}, \quad (4.2)$$

Level 2.

$$u_{\alpha, r} := \arg \min_{u \in BC V_{\alpha}^r(I)} \left\{ \int_I |u - u_0|^2 dx + |u|_{ICTV_{\alpha}^r(I)} \right\}, \quad (4.3)$$

where $u_c \in L^1(I)$ represents a noise-free test picture, and $u_0 \in L^1(I)$ is the noisy image. Note that we only allow the parameters α and the order r of regularizers to lie within a prescribed finite range. This is needed for the numerical realization of our model and also to force the optimal reconstructed image $u_{\bar{\alpha}, \bar{r}}$ to remain inside our proposed space $BCV_{\bar{\alpha}}^{\bar{r}}(I)$ (see Proposition 3.8). In particular, if some of the components of α blow up to ∞ , we might end up in the space $W^{r,1}(I)$, which is outside the purview of this paper. We point out that no upper bound on R is required. Thus, despite the box constraint our analysis still incorporates a large class of image reconstruction regularizers, such as TV and $ICTV^2$ (see, e.g., [19]).

Before we state the main theorem of this section, we prove a technical lemma that will guarantee the existence of a unique solution to (4.3).

Lemma 4.1. *For every $r \in [1, R]$, and $\alpha \in \mathbb{R}_+^{\lfloor r \rfloor + 1}$ there exists a unique $u_{\alpha, r} \in BC V_{\alpha}^r(I)$ solving the minimum problem (4.3).*

Proof. Let $\{u_n\} \subset BC V_{\alpha}^r(I)$ be a minimizing sequence for (4.3). By Lemma 3.7, $\{u_n\}$ is uniformly bounded in $BV(I)$. Thus there exists $\bar{u} \in BV(I)$ such that

$$u_n \rightharpoonup^* \bar{u} \quad \text{weakly* in } BV(I),$$

and hence also strongly in $L^2(I)$. The thesis follows then by Proposition 3.8 and by the strict convexity of the functional. \square

We are now in a position to prove existence and uniqueness of solutions to our learning scheme.

Theorem 4.2. *Let $u_0, u_c \in BV(I)$. Under the box constraint (4.1), the learning scheme (\mathcal{R}) admits a unique solution $(\bar{\alpha}, \bar{r}) \in [A, B]^{\lfloor \bar{r} \rfloor + 1} \times [1, R]$ and provides an associated optimally reconstructed image $u_{\bar{\alpha}, \bar{r}} \in BC V_{\bar{\alpha}}^{\bar{r}}(I)$.*

Proof. Let $\{(\alpha_n, r_n)\}$ be a minimizing sequence for (4.2), with $(\alpha_n, r_n) \in [A, B]^{\lfloor r_n \rfloor + 1} \times [1, R]$ for every $n \in \mathbb{N}$. Let u_{α_n, r_n} be the unique solution to (4.3) provided by Lemma 4.1. By (4.3), there holds

$$|u_{\alpha_n, r_n}|_{BC V_{\alpha_n}^{r_n}(I)} \leq |u_0|_{ICTV_{\alpha_n}^{r_n}(I)} \leq |u'_0|_{\mathcal{M}_b(I)}$$

for every $n \in \mathbb{N}$. There exists $m \in \mathbb{N}$ with $m \in [1, R]$, and $\bar{r} \in [1, R]$ such that, up to the extraction of a (non-relabeled) subsequence, there holds $\lfloor r_n \rfloor \rightarrow m$, and $r_n \rightarrow \bar{r}$. Note that for n big enough we have $\alpha_n \in [A, B]^{m+1}$, and $r_n = m + s_n$, with $s_n \rightarrow \bar{s}$, and $\bar{s} \in [0, 1]$. We distinguish two cases.

Case 1: $\bar{r} \notin \mathbb{N}$. Then $m = \lfloor \bar{r} \rfloor$ and $\bar{s} \in (0, 1)$.

Case 2: $\bar{r} \in \mathbb{N}$. Then either $m = \lfloor \bar{r} \rfloor$ and $\bar{s} = 0$, or $m = \lfloor \bar{r} \rfloor - 1$ and $\bar{s} = 1$.

In both cases Proposition 3.8 yields the existence of a map $u_{\bar{\alpha}, \bar{r}} \in BC V_{\bar{\alpha}}^{\bar{r}}(I)$ such that

$$u_{\alpha_n, r_n} \xrightarrow{*} u_{\bar{\alpha}, \bar{r}} \quad \text{weakly* in } BV(I),$$

thus, in particular, strongly in $L^2(I)$. The existence of solutions follows then by lower semicontinuity, whereas the uniqueness is a direct consequence of the strict convexity of the L^2 -error norm. \square

5. THE NUMERICAL SIMULATION

We present a first-order primal-dual algorithm to solve the minimization problem (4.3). The construction builds upon the general framework of non-smooth convex optimization problems analyzed in [8], and combines it with the numerical techniques for Besov spaces described in [25].

We provide only a sketch of the numerical algorithm, as the numerical implementation of our scheme is not the main focus of this paper. We recall the general framework analyzed in [8]. Let X, Y be two finite-dimensional real vector spaces and let $L: X \rightarrow Y$ be a continuous linear operator. Consider the generic saddle-point problem

$$\min_{x \in X} \max_{y \in Y} \langle Lx, y \rangle + G(x) - F^*(y) \quad (5.1)$$

where $G: X \rightarrow [0, +\infty)$ and $F^*: Y \rightarrow [0, \infty)$ are proper, convex, lower semicontinuous functionals, and where F^* denotes the conjugate of F . Note that this saddle-point problem is a primal-dual formulation of the nonlinear primal problem

$$\min_{x \in X} F(Lx) + G(x)$$

with the corresponding dual problem

$$\max_{y \in Y} -(G^*(-L^*y) + F^*(y)).$$

We define the resolvent operator of F via

$$x = (I + \sigma \partial F)^{-1}(y) = \arg \min_x \left\{ \frac{\|x - y\|^2}{2\sigma} + F(x) \right\}.$$

The resolvent operator for G is defined analogously. In [8] the following scheme for finding the solutions x and y is proposed:

1. Initialization: Choose $\tau, \sigma > 0$, $\theta \in [0, 1]$, $(x^0, y^0) \in X \times Y$ and set $\bar{x}^0 = x^0$.
2. Iterations ($n \geq 0$): Update x^n, y^n, \bar{x}^n as follows:

$$\begin{cases} y^{n+1} &= (I + \sigma \partial F^*)^{-1}(y^n + \sigma L \bar{x}^n) \\ x^{n+1} &= (I + \tau \partial G)^{-1}(x^n - \tau L^* y^{n+1}) \\ \bar{x}^{n+1} &= 2x^{n+1} - x^n. \end{cases}$$

Then the fractional $ICTV^{k+s}$ space proposed in Definition 3.1 can be included into the above framework. In view of Theorem 2.6 we may apply the minimax formulation of Besov

spaces presented in [25]. For $k = 1$ and $0 < s < 1$, we consider the minimax problem

$$\begin{aligned} & \min_{u \in L^1} \frac{1}{2} \|u - u_0\|_{L^2}^2 + |u|_{ICTV_\alpha^{1+s}(I)} \\ &= \min_{u, v_0 \in L^1} \left\{ \frac{1}{2} \|u - u_0\|_{L^2}^2 + \alpha_0 |u' - sv_0|_{\mathcal{M}_b(I)} + \alpha_1 s(1-s) |v_0|_{W^{s, 1+s(1-s)}(I)} \right\} \\ &= \min_{u, v_0} \max_{\varphi, \phi, t} \left\{ \frac{1}{2} \|u - u_0\|_{L^2}^2 + \langle u' - sv_0, \varphi \rangle - I_{\{|\varphi| \leq \alpha_0\}} + \alpha_1 s(1-s) \left(\int_I |K_t^s * v_0|^{s'} \right)^{1/s'} \right\}, \end{aligned}$$

where

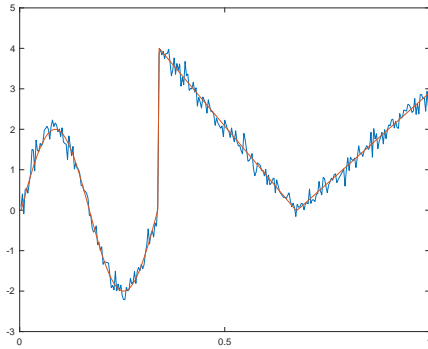
$$I_{\{|\varphi| \leq \alpha_0\}}(x) := \begin{cases} 0 & \text{if } |\varphi(x)| \leq \alpha_0 \\ +\infty & \text{otherwise in } [0, 1], \end{cases}$$

and K_t^s is the operator defined in [25, Section 3]. The reformulation above allows then to identify the operators L , G , and F in (5.1).

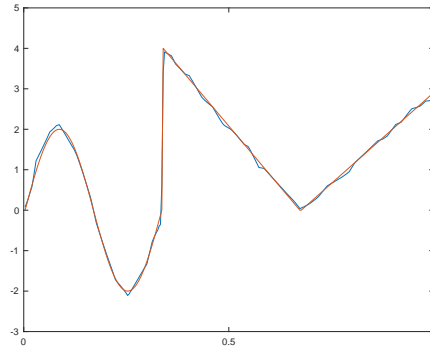
We present in Figure 2 below a simulation of our learning scheme (\mathcal{R}). Choose u_c and u_0 as the blue and red signals in Figure 1A, respectively. Let $A = 0.001$, $B = 2.5$, and $R = 2$. We show in Figure 2 that the L^2 -error function

$$\mathcal{E}(\alpha, r) := \int_I |u_{\alpha, r}(x) - u_c(x)|^2 dx$$

is minimized for $\bar{\alpha}_0 = \bar{\alpha}_1 = 0.17$, and for $\bar{r} = 1.13$.



(A) u_c in red and u_0 in blue.



(B) The best denoised image $u_{\bar{\alpha}, \bar{r}}$ in blue

FIGURE 1. The artificial noise is generated by using a Gaussian noise distribution. In Figure 1B it is also shown that $u_{\bar{\alpha}, \bar{r}}$ mitigates the staircasing effect.

This provides an example in which the optimal reconstruction is achieved for non-integer values of the parameter r .

6. FUTURE WORK

We conclude this paper with an outlook on current and future work, providing a generalization of the learning scheme (\mathcal{R}) to higher dimensions and to a broader setting. Throughout this section $Q := (0, 1) \times (0, 1)$ is the unit square.

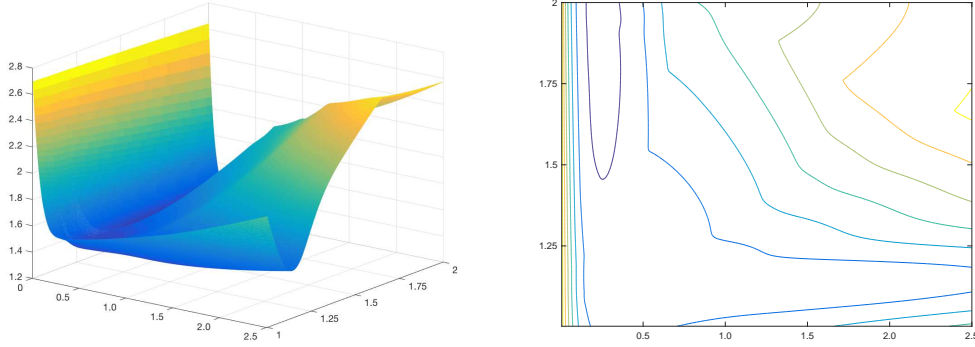
(A) The error function $\mathcal{E}(\alpha, r)$ (B) A plot of the contour lines of $\mathcal{E}(\alpha, r)$

FIGURE 2. $A = 0.001 \leq \alpha_0 = \alpha_1 \leq B = 2.5$, $1 \leq r \leq R = 2$. The error function $\mathcal{E}(\alpha, r)$ attains its minimum at $\bar{\alpha}_0 = 0.2450$ and $\bar{r} = 1.9450$.

6.1. \mathcal{A} - \mathcal{B} Morrey-quasiconvex regularizer. The framework developed in Section 3 leads naturally to consider the \mathcal{A} - \mathcal{B} -quasiconvexity theory (see [13, 24]) to further extend our construction of regularizers. To be precise, let $N = 1, 2$, or $N = 3$, and let \mathcal{A} be a second order differential operator

$$\mathcal{A}u := \sum_{i,j=1,2} A^{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} \quad \text{for every } u \in L^1_{\text{loc}}(Q; \mathbb{M}^{N \times N}),$$

where $A^{ij} \in \mathbb{M}^{\ell \times (N \times N)}$, $\ell \in \mathbb{N}$, $i, j = 1, \dots, N$, and let \mathcal{B} be a first order differential operator defined as

$$\mathcal{B}u := \sum_{k=1,2} B^k \frac{\partial u}{\partial x_k} \quad \text{for every } u \in L^1_{\text{loc}}(Q; \mathbb{R}^N),$$

where $B^k \in \mathbb{R}^{(N \times N) \times N}$ for $k = 1, \dots, N$, and $\mathcal{A}\mathcal{B}u = 0$ for every $u \in L^1_{\text{loc}}(Q; \mathbb{R}^N)$. For every $\alpha, \beta \in \mathbb{R}^+$, and $u \in L^1(\Omega)$ we define the seminorm

$$TV^2_{\mathcal{A}, \mathcal{B}, \alpha, \beta}(u) := \inf \left\{ \alpha \|\nabla u - p\|_{\mathcal{M}_b(Q; \mathbb{R}^N)} + \beta \int_Q \mathcal{Q}_{\mathcal{A}} f(\mathcal{B}p(x)) dx : \right. \\ \left. p \in L^1(Q; \mathbb{R}^N), \mathcal{B}p \in L^1(Q; \mathbb{M}^{N \times N}) \right\},$$

where $\mathcal{Q}_{\mathcal{A}} f$ represents the \mathcal{A} -quasiconvex envelope of f , the function $f : \mathbb{M}^{N \times N} \rightarrow [0, +\infty)$ is Lipschitz continuous, and there exists $C > 0$ such that $C^{-1}|\xi| \leq f(\xi) \leq C|\xi|$ for every $\xi \in \mathbb{M}^{N \times N}$. Note that when $f(\cdot) = |\cdot|$, $\mathcal{B} = \text{sym} \nabla$, and $\mathcal{A} = \text{curl curl}$, we have $TV^2_{\mathcal{A}, \mathcal{B}, \alpha, \beta}(u) = TGV^2_{\alpha, \beta}(u)$, where $TGV^2_{\alpha, \beta}(u)$ is defined in [40]. The analysis of the following learning scheme with respect to $TV^2_{\mathcal{A}, \mathcal{B}, \alpha, \beta}(u)$ is the subject of [16]:

Level 1.

$$(\bar{\mathcal{A}}, \bar{\mathcal{B}}, \bar{\alpha}, \bar{\beta}) := \arg \min \left\{ \int_Q |u_{\mathcal{A}, \mathcal{B}, \alpha, \beta} - u_c|^2 dx, \right. \\ \left. A^{ij} \in \mathbb{M}^{\ell \times (N \times N)}, B^k \in \mathbb{M}^{(N \times N) \times N}, i, j, k = 1, \dots, N \right\},$$

Level 2.

$$u_{\mathcal{A},\mathcal{B},\alpha,\beta} := \arg \min_{u \in L^1(Q)} \left\{ \frac{1}{2} \int_Q |u - u_c|^2 dx + TV_{\mathcal{A},\mathcal{B},\alpha,\beta}^2(u) \right\}.$$

6.2. The spatially dependent learning scheme. In [28], an improved version of scheme (B) is proposed. For $K \in \mathbb{N}$, $Q_K \subset \mathbb{R}^2$ denotes a cube with its faces normal to the orthonormal basis of \mathbb{R}^2 , and with side-length greater than or equal to $1/K$. Define \mathcal{P}_K to be a collection of finitely many Q_K such that

$$\mathcal{P}_K := \left\{ Q_K \subset \mathbb{R}^N : Q_K \text{ are mutually disjoint, } Q \subset \bigcup Q_K \right\},$$

and

$$\mathcal{V}_K := \{\text{all possible } \mathcal{P}_K\}. \quad (6.1)$$

For $K = 0$ we set $Q_0 := Q$, hence $\mathcal{P}_0 = \{Q\}$. We propose a trilevel learning scheme (P) as follows:

Level 1.

$$u_{\mathcal{P}} := \arg \min_{K \geq 0, \mathcal{P}_K \in \mathcal{V}_K} \left\{ \int_Q |u_c - u_{\mathcal{P}_K}|^2 dx \right\}$$

Level 2. for $x \in Q_K \in \mathcal{P}_K$,

$$u_{\mathcal{P}_K}(x) := \arg \min_{u \in SBV(Q_K)} \left\{ \frac{1}{2} \int_{Q_K} |u - u_0|^2 dx + \alpha_{Q_K} TV(u) \right\},$$

Level 3.

$$\begin{aligned} \alpha_{Q_K} &:= \arg \min_{\alpha > 0} \int_{Q_K} |u_{\alpha,K} - u_c|^2 dx \\ u_{\alpha,K} &:= \arg \min_{u \in SBV(Q_K)} \left\{ \frac{1}{2} \int_{Q_K} |u - u_0|^2 dx + \alpha TV(u) \right\}. \end{aligned}$$

The learning Level 3 determines the optimal parameter α in each subdomain Q_K by utilizing scheme (B) in each Q_K , and the learning Level 2 glues the optimal reconstructed images produced in each subdomain, while Level 1 searches for the best combination of those subdomains. Scheme (P) allows us to perform the denoising procedure “pointwisely”, and it is an improvement of (1.2) since \mathcal{V}_K defined in (6.1) is nested.

Theorem 6.1 ([28]). *Assume that the noise η has locally zero average, that is,*

$$\int_{Q_K} \eta = 0 \text{ for any } Q_K \in \mathcal{P}_K. \quad (6.2)$$

Define, for each $K \in \mathbb{N}$,

$$\mathcal{P}(K) := \min \left\{ \int_Q |u_c - u_{\mathcal{P}_K}|^2 dx, \mathcal{P}_K \in \mathcal{V}_K, K \geq 0 \right\}.$$

Then $\lim_{K \rightarrow \infty} \mathcal{P}(K) = 0$.

Condition (6.2) is rather strong and provides a restriction on the range of applicability of Theorem 6.1. With our new learning scheme at hand, we are working toward a relaxation of (6.2), in which scheme (B) in Level 3 in (P) would be replaced with our new learning scheme (R).

ACKNOWLEDGEMENTS

P. Liu is partially funded by the National Science Foundation under Grant No. DMS - 1411646. E. Davoli is supported by the Austrian Science Fund (FWF) project P27052. The authors wish to thank Irene Fonseca for suggesting this project and for many helpful discussions and comments. The authors also thank Giovanni Leoni for comments on the subject of Section 2.

REFERENCES

- [1] R.A. Adams, J.J.F. Fournier. *Sobolev spaces*. Elsevier/Academic Press, Amsterdam, second edition, 2003.
- [2] L. Ambrosio, N. Fusco, D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.
- [3] M. Bergounioux. Optimal control of problems governed by abstract elliptic variational inequalities with state constraints. *SIAM J. Control Optim.* **36**(1) (1998), 273–289.
- [4] J. Bourgain, H. Brezis, P. Mironescu. Limiting embedding theorems for $W^{s,p}$ when $s \uparrow 1$ and applications. *J. Anal. Math.* **87** (2002), 77–101.
- [5] J. Bourgain, H. Brezis, P. Mironescu. Another look at sobolev spaces. *Optimal Control and Partial Differential Equations*, J. L. Menaldi, E. Rofman, A. Sulem (Eds.), a volume in honor of A. Bensoussan’s 60th birthday, IOS Press, Amsterdam, pages 439–455, 2001.
- [6] K. Bredies, T. Valkonen. Inverse problems with second-order total generalized variation constraints. *Proceedings of SampTA 2011 - 9th International Conference on Sampling Theory and Applications*, Singapore, 2011.
- [7] M. Burger, K. Papafitsoros, E. Papoutsellis, C.-B. Schönlieb. Infimal convolution regularisation functionals of BV and L^p spaces. Part I: the finite p case. *J. Math. Imaging Vision*, **55**(3) (2016), 343–369.
- [8] A. Chambolle, T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision*, **40**(1) (2011), 120–145.
- [9] T. Chan, A. Marquina, P. Mulet. High-order total variation-based image restoration. *SIAM J. Sci. Comput.* **22**(2) (2000), 503–516.
- [10] D. Chen, S. Sun, C. Zhang, Y. Chen, D. Xue. Fractional-order $TV-L^2$ model for image denoising. *Central European Journal of Physics*, **11**(10) (2013), 1414–1422.
- [11] Y. Chen, T. Pock, R. Ranftl, H. Bischof. Revisiting loss-specific training of filter-based mrfs for image restoration. In J. Weickert, M. Hein, and B. Schiele, editors, *Pattern Recognition: 35th German Conference, GCPR 2013, Saarbrücken, Germany, September 3-6, 2013. Proceedings*, pages 271–281. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [12] Y. Chen, R. Ranftl, T. Pock. Insights into analysis operator learning: From patch-based sparse models to higher order mrfs. *IEEE Transactions on Image Processing*, **23**(3) (2014), 1060–1072.
- [13] B. Dacorogna. *Direct methods in the calculus of variations*. Springer, New York, second edition, 2008.
- [14] G. Dal Maso, I. Fonseca, G. Leoni, M. Morini. A higher order model for image restoration: the one-dimensional case. *SIAM J. Math. Anal.* **40**(6) (2009), 2351–2391.
- [15] J. Dávila. On an open question about functions of bounded variation. *Calc. Var. Partial Differential Equations*, **15**(4) (2002), 519–527.
- [16] E. Davoli, I. Fonseca, P. Liu. Image reconstruction and \mathcal{A} - \mathcal{B} quasiconvexity. *In preparation*.
- [17] E. Davoli, P. Liu. Optimal parameters and regularizers for image reconstruction problems - two dimensional setting and numerical analysis. *In preparation*.
- [18] J. C. De los Reyes, C.-B. Schönlieb. Image denoising: learning the noise model via nonsmooth PDE-constrained optimization. *Inverse Probl. Imaging*, **7**(4) (2013), 1183–1214.
- [19] J. C. De los Reyes, C.-B. Schönlieb, T. Valkonen. Bilevel parameter learning for higher-order total variation regularisation models. *J. Math. Imaging Vision*, (2016), 1–25.
- [20] J.C. De los Reyes, C.-B. Schönlieb, T. Valkonen. The structure of optimal parameters for image restoration problems. *J. Math. Anal. Appl.* **434**(1) (2016), 464–500.
- [21] E. Di Nezza, G. Palatucci, E. Valdinoci. Hitchhiker’s guide to the fractional Sobolev spaces. *Bull. Sci. Math.* **136**(5) (2012), 521–573.
- [22] J. Domke. Generic methods for optimization-based modeling. In N.D. Lawrence and M.A. Girolami, editors, *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics (AISTATS-12)*, volume 22, pages 318–326, 2012.

- [23] J. Domke. Learning graphical model parameters with approximate marginal inference. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**(10) (2013), 2454–2467.
- [24] I. Fonseca and S. Müller. \mathcal{A} -quasiconvexity, lower semicontinuity, and Young measures. *SIAM J. Math. Anal.* **30**(6) (1999), 1355–1390.
- [25] J.B. Garnett, T.M. Le, Y. Meyer, L.A. Vese. Image decompositions using bounded variation and generalized homogeneous Besov spaces. *Appl. Comput. Harmon. Anal.* **23**(1) (2007), 25–56.
- [26] N.S. Landkof. *Foundations of modern potential theory*. Springer-Verlag, New York-Heidelberg, 1972.
- [27] G. Leonori. *A first course in Sobolev spaces*. American Mathematical Society, Providence, RI, 2009.
- [28] P. Liu. The improved bi-level learning scheme for TV . *In preparation*.
- [29] V.G. Maz'ya, T.O. Shaposhnikova. On the Bourgain, Brezis, and Mironescu theorem concerning limiting embeddings of fractional Sobolev spaces. *J. Funct. Anal.* **195**(2) (2002), 230–238.
- [30] V.G. Maz'ya, T.O. Shaposhnikova. *Theory of Sobolev multipliers*. Springer-Verlag, Berlin, 2009.
- [31] K. Papafitsoros, K. Bredies. A study of the one dimensional total generalised variation regularisation problem. *Inverse Probl. Imaging*, **9**(2) (2015), 511–550.
- [32] K. Papafitsoros, C.B. Schönlieb. A combined first and second order variational approach for image reconstruction. *J. Math. Imaging Vision*, **48**(2) (2014), 308–338.
- [33] L.I. Rudin. Segmentation and restoration using local constraints. *Cognitech Report* **36** (1993).
- [34] L.I. Rudin, S. Osher. Total variation based image restoration with free local constraints. In *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 1, pages 31–35 vol.1, Nov 1994.
- [35] L.I. Rudin, S. Osher, E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, **60** (1992), 259–268.
- [36] J. Simon. Sobolev, Besov and Nikol'skiĭ fractional spaces: embeddings and comparisons for vector valued spaces on an interval. *Ann. Mat. Pura Appl.* **157** (1990), 117–148.
- [37] M.F. Tappen. Utilizing variational optimization to learn markov random fields. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2007.
- [38] M.F. Tappen, C. Liu, E.H. Adelson, W.T. Freeman. Learning gaussian conditional random fields for low-level vision. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2007.
- [39] H. Triebel. *Theory of function spaces*. Modern Birkhäuser Classics. Birkhäuser/Springer Basel AG, Basel, 2010.
- [40] T. Valkonen, K. Bredies, F. Knoll. Total generalized variation in diffusion tensor imaging. *SIAM J. Imaging Sci.* **6**(1) (2013), 487–525.